



ECOLE NATIONALE SUPÉRIEUR PARIS-SACLAY  
MASTER MATHÉMATIQUES VISION APPRENTISSAGE

# TP 2: Détection de Communautés dans les Graphes

OREISTEIN Pierre

COURS: THÉORIE DES MATRICES ALÉATOIRES

APRIL 12, 2019

# 1 - Observations préliminaires:

## 1.1

Dans la suite, les calculs sont fait en considérant les  $(q_i)_{i=1\dots n}$  connues. Posons:

$$M = \begin{pmatrix} M_{11} & \cdots & M_{1K} \\ \vdots & \ddots & \vdots \\ M_{K1} & \cdots & M_{KK} \end{pmatrix} \in \mathcal{M}_{K \times K}(\mathbb{R}), \quad \mathbf{1} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \quad \text{et} \quad C = \mathbf{1} + \frac{1}{\sqrt{n}} M \in \mathcal{M}_{K \times K}(\mathbb{R})$$

De même posons  $\forall k \in \llbracket 1, K \rrbracket, \quad n_k = |C_k|$  et:

$$J = (j_1, \dots, j_K) \in \mathcal{M}_{n \times K}(\mathbb{R}), \quad \text{avec} \quad j_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{1}_{n_k} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{et} \quad q = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}$$

On peut alors remarquer le résultat suivant:

$$JCJ^T = \begin{pmatrix} C_{11} & \cdots & C_{11} & \cdots & C_{1k} & \cdots & C_{1k} & \cdots & C_{1K} & \cdots & C_{1K} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ C_{11} & \cdots & C_{11} & \cdots & C_{1k} & \cdots & C_{1k} & \cdots & C_{1K} & \cdots & C_{1K} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_{k1} & \cdots & C_{k1} & \cdots & C_{kk} & \cdots & C_{kk} & \cdots & C_{kK} & \cdots & C_{kK} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ C_{k1} & \cdots & C_{k1} & \cdots & C_{kk} & \cdots & C_{kk} & \cdots & C_{kK} & \cdots & C_{kK} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_{K1} & \cdots & C_{K1} & \cdots & C_{Kk} & \cdots & C_{Kk} & \cdots & C_{KK} & \cdots & C_{KK} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ C_{K1} & \cdots & C_{K1} & \cdots & C_{Kk} & \cdots & C_{Kk} & \cdots & C_{KK} & \cdots & C_{KK} \end{pmatrix}$$

On alors une matrice par bloc où chaque bloc de  $C_{ij}\mathbf{1}$  est une matrice de taille  $n_i \times n_j$ .

De plus, soient  $i, j \in \llbracket 1, n \rrbracket$ , on a:

$$A_{ij} = \mathbb{E}[A_{ij}] + X_{ij} \quad \text{avec} \quad X_{ij} + q_i q_j C_{ab} \sim \text{Ber}(q_i q_j C_{ab})$$

$$A_{ij} = q_i q_j C_{ab} + X_{ij}$$

On peut alors remarquer que:

- La matrice  $X$  définit par les  $(X_{ij})_{i,j=1\dots n}$  est à entrées indépendantes car les  $X_{ij}$  dépendent de loi de Bernoulli indépendantes.
- $\mathbb{E}[X] = (1 - q_i q_j C_{ab}) q_i q_j C_{ab} + (-q_i q_j C_{ab})(1 - q_i q_j C_{ab}) = 0$

Enfin, si on pose  $Q = \text{diag}(q_i)$  et quitte à ordonner les noeuds du graphe selon leur classe, on obtient alors le résultat suivant:

$$A = Q(JCJ^T)Q^T + X$$

Où:

- $C$  est une matrice de rang au plus  $K$ .
- $X$  est une matrice à entrées aléatoires indépendantes de moyenne nulle.

## 1.2

Avant tout calcul, on peut:

- Décomposer  $C$  par :  $\mathbf{1} + \frac{1}{\sqrt{n}}M$
- Remarquer que  $qq^T = Q\mathbf{1}Q$

En utilisant ces deux résultats, on obtient donc:

$$\begin{aligned} B &= A - qq^T \\ &= QJ\mathbf{1}J^TQ^T + \frac{1}{\sqrt{n}}qJMJ^Tq^T + X - Q\mathbf{1}Q^T \\ \text{On peut alors remarquer que: } \mathbf{1}_{K \times n} &= \mathbf{1}_{K \times K}J^T \quad \text{et} \quad \mathbf{1}_{n \times n} = J\mathbf{1}_{K \times n} \\ &= \frac{1}{\sqrt{n}}QJMJ^TQ^T + X \end{aligned}$$

Or  $M$  est une matrice de rang au plus  $K$  et  $X$  est inchangé. On obtient donc les mêmes propriétés sur  $B$  que sur  $A$ .

## 1.3

Dans cette section, on se place dans le cas où  $K = 3$ . On va s'intéresser au spectre de  $B$  pour différents choix des valeurs  $(q_i)_{i=1\dots n}$  et pour la génération de la matrice  $M$ . En pratique, nous avons implémentés différentes méthodes pour générer ces valeurs:

- *Pour la génération des  $(q_i)_{i=1\dots n}$ :*
  - *Constant:* Même valeur de  $q_i$  pour tout  $i$ . On teste pour  $q_i \in \{0.1, 0.5, 0.9\}$
  - *Uniforme:* On génère les  $q_i$  de manière uniforme autour d'une valeur centrale  $q_0$ . En pratique, on a testé pour  $q_0 \in \{0.1, 0.5, 0.9\}$  avec des valeurs prises dans un intervalle de taille 0.5.
  - *Deux valeurs:* On génère les  $q_i$  en tirant de manière équiprobable entre deux valeurs  $\{q^{(1)}, q^{(2)}\}$ . En pratique, on a testé les couples de valeurs suivants:  $(0.3, 0.7)$ ,  $(0.45, 0.9)$  et  $(0.1, 0.9)$
- *Pour la génération de  $M$ :*
  - *Identité:*  $M$  est égale à la matrice identité:  $I_{K \times K}$
  - *Faible corrélation inter-groupe:* Ici,  $M = 5 * I_{K \times K} + Y$  où  $Y$  est une matrice dont les entrées sont des valeurs entières comprises 0 et 1. Puis on symétrise  $M$  via  $M = \frac{1}{2}(M + M^T)$ .
  - *Forte Corrélation inter-groupe:* Ici,  $M = 5 * I_{K \times K} + Y$  où  $Y$  est une matrice dont les entrées sont des valeurs entières comprises 0 et 4. Puis on symétrise  $M$  via  $M = \frac{1}{2}(M + M^T)$ .

Grâce à ces méthodes, nous obtenons les résultats suivants:

- $q$  est constant:

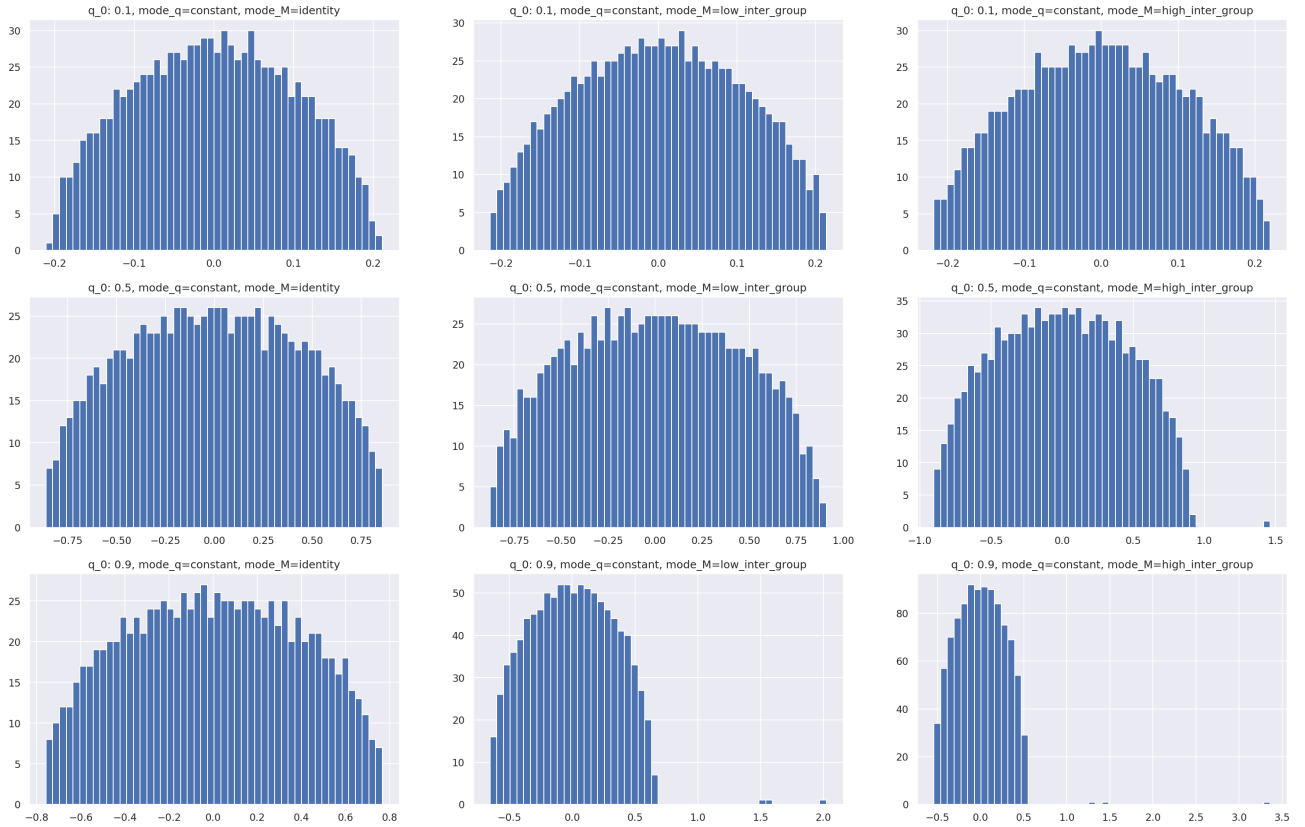


Figure 1 – Distribution des valeurs propres du spectre de  $\frac{1}{\sqrt{n}}B$  pour  $n = 1000$  dans le cas où les  $q_i$  sont constants. Sur la première ligne, les  $q_i$  sont égaux à 0.1, sur la deuxième à 0.5 et sur la dernière 0.9. Sur la première colonne,  $M$  est généré via la méthode *Identité*, sur la seconde, via la méthode *Faible corrélation inter-groupe* et sur la dernière *Forte corrélation inter-groupe*.

On observe que:

- Plus la valeur constante de  $q_i$  augmente, plus on observe de valeurs propres isolées. En particulier on peut observer (peu visible du fait de l'échelle des graphes) 3 valeurs propres isolées sur les deux dernières figures de la dernière ligne (lorsque  $q_i$  est maximal).
- Plus la corrélation entre les différents groupes est importante, plus la probabilité d'observer des valeurs propres isolées est importante (e.g. dernière colonne pour laquelle  $M$  est générée avec une corrélation importante entre tous ses termes).
- Un spectre qui se disperse selon la loi du demi-cercle. Ceci est logique car lorsque  $q$  est constant,  $\frac{1}{\sqrt{n}}X_n$  correspond, à renormalisation près, à une matrice de Wigner. Dès lors il est logique d'observer une distribution en demi-cercle du fait du théorème de Wigner.

- $q$  est uniforme:

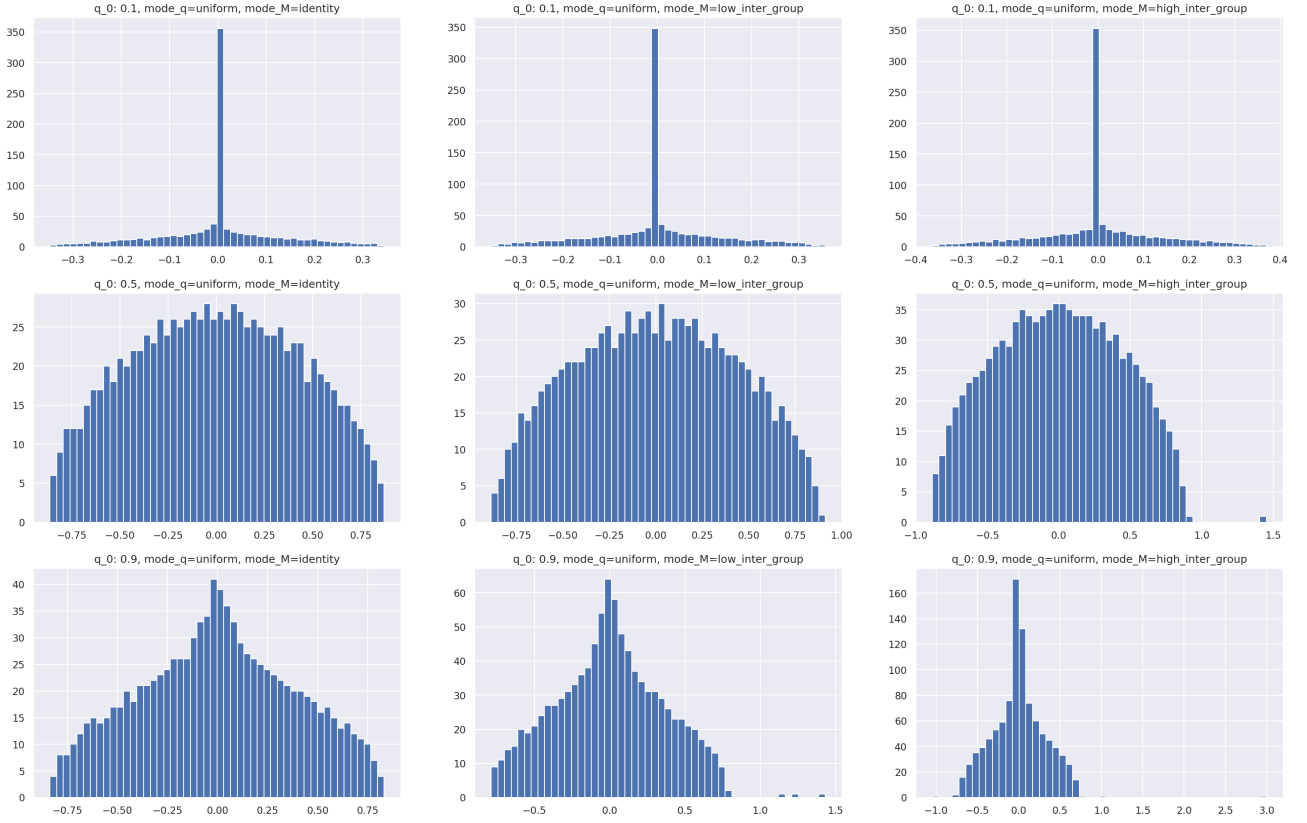


Figure 2 – Distribution des valeurs propres du spectre de  $\frac{1}{\sqrt{n}}B$  pour  $n = 1000$  dans le cas où les  $q_i$  sont générées de manière uniforme autour d’une valeur centrale. Sur la première ligne, la valeur moyenne des  $q_i$  vaut 0.1, sur la deuxième à 0.5 et sur la dernière 0.9. Sur la première colonne,  $M$  est généré via la méthode *Identité*, sur la seconde, via la méthode *Faible corrélation inter-groupe* et sur la dernière *Forte corrélation inter-groupe*.

De la même manière que précédemment, on observe que:

- Plus la valeur constante de  $q_0$  augmente, plus on observe de valeurs propres isolées. En particulier on peut observer 3 valeurs propres isolées sur les deux dernières figures de la dernière ligne (lorsque  $q_0$  est maximal et vaut 0.9).
- Plus la corrélation entre les différents groupes est importante, plus la probabilité d’observer des valeurs propres isolées est importante (e.g. dernière colonne pour laquelle  $M$  est générée avec une corrélation importante entre tous ses termes).
- La distribution du spectre est différente de précédemment avec une concentration plus forte autour de la valeur propre 0 en particulier dans le cas des valeurs de  $q_0$  extrêmes. Dans le cas de la première ligne, comme on génère des  $q_i$  de manière uniforme entre  $-0.15$  et  $0.35$  puis on prend la partie positive. On a donc une valeur moyenne des  $q_i$  qui est très proche de 0. Dès lors il est logique d’observer un phénomène de saturation autour de 0 pour les valeurs propres de  $X$ . Dans le cas de la dernière ligne, on génère de manière uniforme des  $q_i$  entre 0.65 et 1.15 puis on prend le minimum avec 1. Dès lors, les  $q_i$  ont une valeur moyenne proche de 1. Ainsi, les  $X_{ij}$  prennent très souvent la valeur 1 et  $X$  est alors de rang faible, d’où une concentration des valeurs propres autour de 0. Dans le cas  $q_0 = 0.5$  (ligne du milieu), on retrouve une distribution en demi-cercle ce qui est logique. En effet, cette fois, il n’y a pas de phénomène de saturation car les  $q_i$  prennent des valeurs moyennes dans l’intervalle  $[0, 1]$  et on retombe dès lors sur des matrices de Wigner pour  $\frac{1}{\sqrt{n}}X_n$ .

- $q_i$  peut prendre deux valeurs possibles:

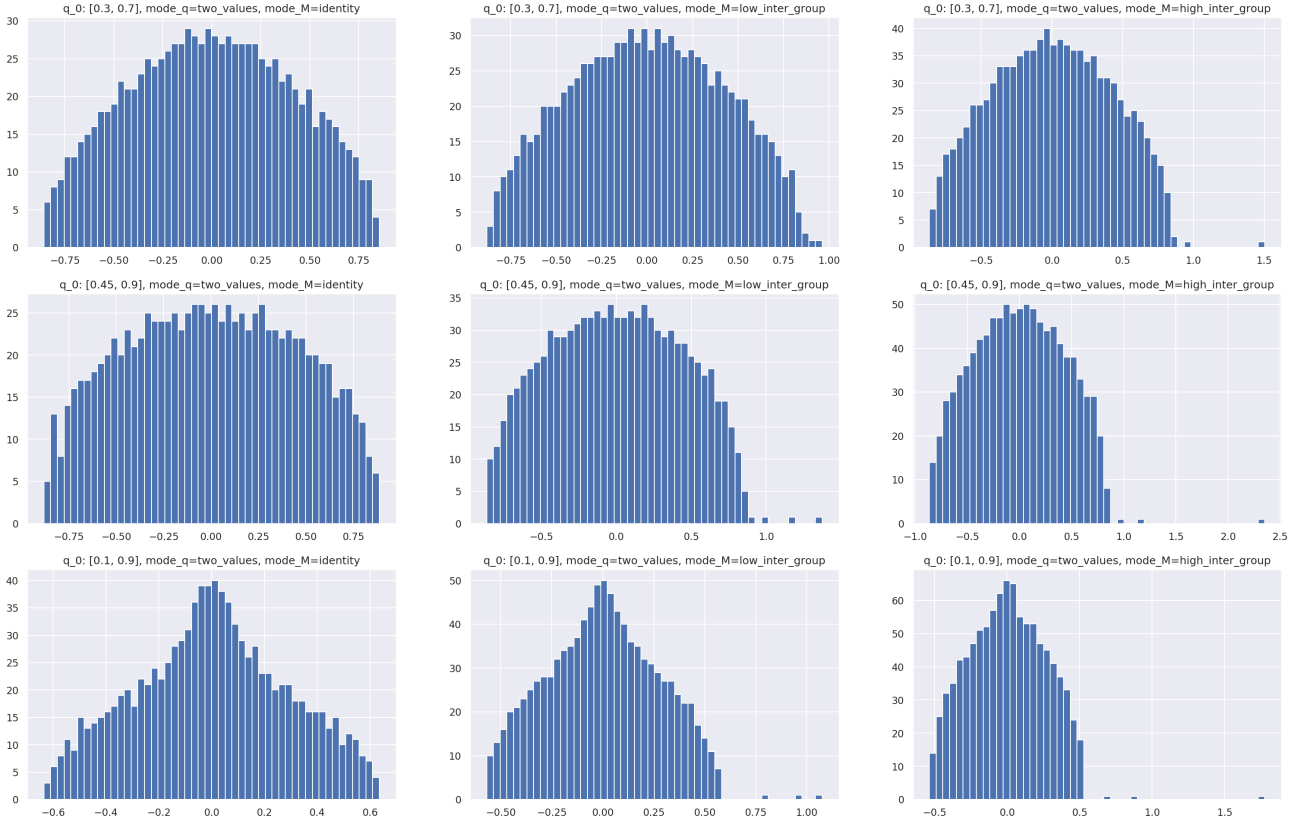


Figure 3 – Distribution des valeurs propres du spectre de  $\frac{1}{\sqrt{n}}B$  pour  $n = 1000$  dans le cas où les  $q_i$  sont générées à partir de deux valeurs équiprobables. Sur la première ligne, le couple de valeurs possibles  $(q^{(1)}, q^{(2)})$  des  $q_i$  vaut  $(0.3, 0.7)$ , sur la deuxième à  $(0.45, 0.9)$  et sur la dernière  $(0.1, 0.9)$ . Sur la première colonne,  $M$  est généré via la méthode *Identité*, sur la seconde, via la méthode *Faible corrélation inter-groupe* et sur la dernière *Forte corrélation inter-groupe*.

De la même manière que précédemment, on observe:

- Plus la valeur maximale du couple  $(q^{(1)}, q^{(2)})$  augmente, plus on observe de valeurs propres isolées. En particulier on peut observer 3 valeurs propres isolées sur les deux dernières figures de la dernière ligne (lorsque  $q^{(2)}$  est maximal et vaut 0.9).
- Plus la corrélation entre les différents groupes est importante, plus la probabilité d’observer des valeurs propres isolées est importante (e.g. dernière colonne pour laquelle  $M$  est générée avec une corrélation importante entre tous ses termes).
- La distribution du spectre tend vers la loi du demi-cercle dans le cas des deux premières lignes car les  $q_i$  prennent des valeurs plutôt moyennes dans l’intervalle  $[0, 1]$  ce qui n’est pas le cas de la dernière lignes où les  $q_i$  prennent plutôt des valeurs extrêmes.

## 1.4

Dans cette section, on effectue une ACP sur la matrice  $B$  et on représente les données dans l'espace des trois vecteurs principaux dans les cas où l'on observait des spikes. En particulier, on effectue une ACP dans le cas  $n = 1000$ ,  $M$  est générée par la méthode *Forte corrélation inter-groupe* et les  $q_i$  sont constants et prennent la valeur 0.9. Avec ces paramètres, on obtient le graphique ci-dessous :

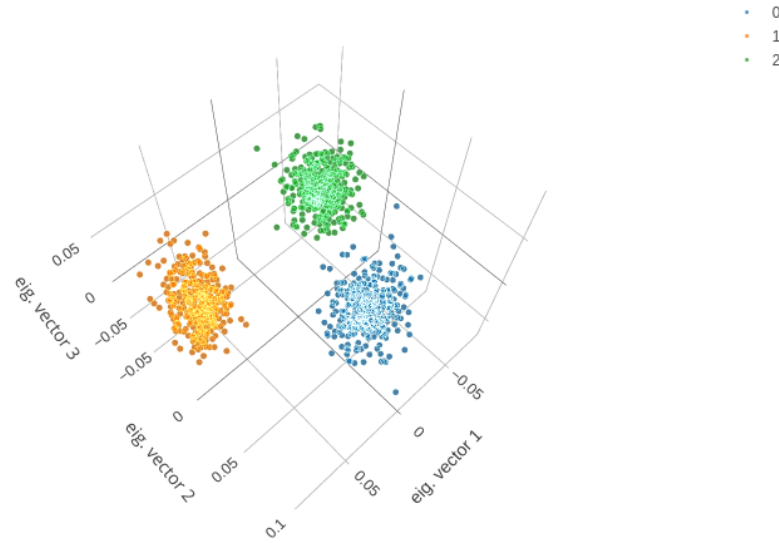


Figure 4 – Représentation des vecteur de la matrice  $\frac{1}{\sqrt{n}}B$  dans l'espace des trois plus grandes valeurs propres avec leur cluster associé

On constate que les trois clusters sont bien séparés. Ainsi, en combinant cette approche avec un algorithme de type *K-means* ou *EM*, on serait capable de faire du clustering de communautés.

Lorsque les  $q_i$  ne sont plus égaux, la matrice  $A$  sera en quelque-sort "bruité" par rapport à  $C$ . Dès lors, l'information  $C$  contenu dans  $A$  sera plus difficilement accessible. C'est précisément ce que l'on observe sur la figure ci-dessous où l'on a à nouveau réalisé une ACP de la matrice  $\frac{1}{\sqrt{n}}B$  avec des  $q_i$  générés selon la méthode *deux valeurs* pour un couple de valeurs égale à (0.1, 0.9) avec la même méthode de génération de  $M$ .

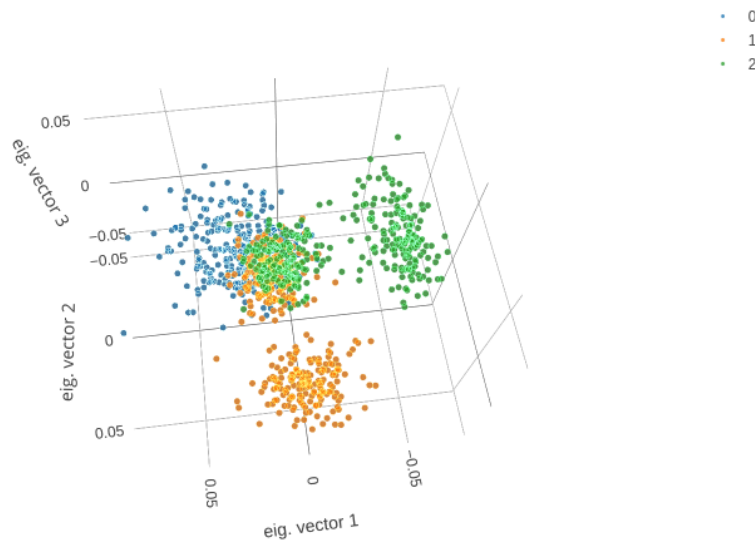


Figure 5 – Représentation des vecteur de la matrice  $\frac{1}{\sqrt{n}}B$  dans l'espace des trois plus grandes valeurs propres avec leur cluster associé

Cette fois-ci on constate qu'une partie des données ne sont plus aisément séparables dans cet espace.

## 2 - Cas Homogène:

Dans cette partie, on considère le cas où les  $q_i$  sont tous égaux et valent  $q_0 \in (0, 1)$ . De même, on se place dans le cas où  $M$  est diagonale.

### 2.1

En reprenant le résultat de la question 1.2, on a:

$$\frac{1}{\sqrt{n}}B = \frac{1}{n}QJM J^T Q^T + \frac{1}{\sqrt{n}}X_n, \quad \text{avec } X_{ij} + q_i q_j C_{ab} \sim \text{Ber}(q_i q_j C_{ab})$$

Ainsi, en utilisant les nouvelles hypothèses sur les  $q_i$  et  $M$ , on en déduit que:

- $Q = q_0 I_n$
- $X_n$  a toujours des entrées de moyenne nulle.

- Pour rappel  $C_{ab} = 1 + \frac{M_{ab}}{\sqrt{n}}$ .

Ainsi, la variance des entrées de  $\frac{1}{\sqrt{n}}X_n$  vaut donc: 
$$\begin{cases} \frac{q_0^2(1-q_0^2)}{n} & , \text{ si } a \neq b \\ \frac{q_0^2}{n} \left(1 + \frac{M_{aa}}{\sqrt{n}}\right) (1 - q_0^2 \left(1 + \frac{M_{aa}}{\sqrt{n}}\right)) & , \text{ sinon } \end{cases}$$

et donc la variance de  $\frac{1}{\sqrt{n}}X_n$  tend asymptotiquement vers  $\frac{q_0^2(1-q_0^2)}{n}$

- Enfin, vu que les entrées de  $\frac{1}{\sqrt{n}}X_n$  suivent une loi de Bernoulli, on en déduit que leur moment d'ordre 4 est fini et décroît avec  $n$ .

Grâce à ces remarques, on va alors pouvoir utiliser le théorème de Wigner:

#### **Theorem 2.1** (Théorème de Wigner)

Soit  $X_n \in \mathbb{R}^{n \times n}$  une matrice carrée à entrées indépendantes, de moyenne nulle, de variance convergent uniformément vers  $\frac{1}{n}$  et de moment d'ordre 4 fini et indépendant de  $n$ . Alors, la mesure spectrale empirique  $\frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i(X_n)}$  converge vers la loi du demi-cercle de mesure  $\mathbb{P}_{sc}$  de densité:

$$\mathbb{P}_{sc}(dx) = \frac{1}{2\pi} \sqrt{(4-x^2)^+} dx$$

On déduit du théorème que lorsque  $n$  tend vers l'infini, la mesure empirique de  $\frac{1}{\sqrt{n}\sqrt{q_0^2(1-q_0^2)}}X_n$  converge vers la loi du demi-cercle de mesure  $\mathbb{P}_{sc}$ .

Si on reprend maintenant l'étude de  $\frac{1}{\sqrt{n}}B$ , on peut interpréter sa décomposition comme la perturbation de la matrice  $\frac{1}{\sqrt{n}}X_n$  par une matrice de rang au plus  $K$ .

Dés lors, on peut suivre la procédure du cours: On cherche  $\lambda \in \mathbb{R}$  telle que:

$$\det\left(-\lambda I_n + \frac{1}{\sqrt{n}}X_n\right) \neq 0 \quad \text{et} \quad \det\left(-\lambda I_n + \frac{1}{\sqrt{n}}X_n + \frac{1}{n}QJM J^T Q^T\right) = 0$$



On peut alors résoudre la deuxième équation:

$$\begin{aligned}
\det \left( -\lambda I_n + \frac{1}{\sqrt{n}} X_n + \frac{1}{n} Q J M J^T Q^T \right) &= 0 \iff \det \left( -\lambda I_n + \frac{1}{\sqrt{n}} X_n + \frac{q_0^2}{n} J M J^T \right) = 0 \\
&\iff \det \left( I_n + \frac{q_0^2}{n} \left( \frac{1}{\sqrt{n}} X_n - \lambda I_n \right)^{-1} J M J^T \right) = 0 \\
\text{Or } Q \frac{1}{\sqrt{n}} X_n (\lambda) &= \left( \frac{1}{\sqrt{n}} X_n - \lambda I_n \right)^{-1} \\
&\iff \det \left( I_n + \frac{q_0^2}{n} Q \frac{1}{\sqrt{n}} X_n (\lambda) J M J^T \right) = 0 \\
&\iff \det \left( I_K + \frac{q_0^2}{n} J^T Q \frac{1}{\sqrt{n}} X_n (\lambda) J M \right) = 0
\end{aligned}$$

Or comme on l'a vu, la mesure spectrale empirique de  $\frac{1}{\sqrt{n}} X_n$  converge p.s. vers la loi du demi-cercle de mesure  $q_0 \sqrt{1 - q_0^2} \mathbb{P}_{sc}(dx)$ . Ainsi  $q_0 \sqrt{1 - q_0^2} g_{sc}$  est la transformée de Stieljes de  $q_0 \sqrt{1 - q_0^2} \mathbb{P}_{sc}$  et l'on a:

$$\forall u^*, v, \quad u^* Q(\lambda) v - g_{sc}(\lambda) (q_0 \sqrt{1 - q_0^2}) u^* v \xrightarrow{n \rightarrow \infty} 0$$

avec  $g_{sc}$  la transformée de Stieljes de  $\mathbb{P}_{sc}$ .

Ainsi, p.s., on a:

$$\det \left( -\lambda I_n + \frac{1}{\sqrt{n}} X_n + \frac{1}{n} Q J M J^T Q^T \right) = 0 \xrightarrow{n \rightarrow \infty} \det \left( I_K + \frac{q_0^3 \sqrt{1 - q_0^2}}{n} g_{sc}(\lambda) J^T J M \right) = 0$$

Or, il se trouve que  $J^T J = \begin{pmatrix} |C_1| & 0 & \dots & 0 \\ 0 & |C_2| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |C_K| \end{pmatrix} \in \mathbb{R}^{K \times K}$

Dès lors on en déduit qu'asymptotiquement, il existe des valeurs propres isolées, si et seulement si:

$$\exists \lambda \in \mathbb{R}, k \in \llbracket 1, K \rrbracket, \quad \text{tel que} \quad c_k M_{kk} = \frac{-1}{q_0^3 \sqrt{1 - q_0^2} g_{sc}(\lambda)}$$

### 2.1.1

On sait que  $g_{sc}(z)$  est l'unique solution de l'équation:

$$g = -\frac{1}{z + g}$$

On spécialise alors l'équation précédente en  $z = \lambda$  et en utilisant le résultat de la question précédente, on en déduit:

$$\begin{aligned}
(g_{sc}(\lambda))^2 + \lambda g_{sc}(\lambda) &= -1 \\
\frac{1}{(q_0^3 \sqrt{1 - q_0^2} c_k M_{kk})^2} + \frac{-\lambda}{q_0^3 \sqrt{1 - q_0^2} c_k M_{kk}} &= -1 \\
\lambda &= q_0^3 \sqrt{1 - q_0^2} c_k M_{kk} \left( 1 + \frac{1}{(q_0^3 \sqrt{1 - q_0^2} c_k M_{kk})^2} \right)
\end{aligned}$$

Malgré nos différents essais, nous ne sommes pas parvenus à reproduire ces résultats via des simulations.

## 2.2

Dans un premier temps, on peut utiliser l'identité de Woodbury. Pour rappel:

**Lemma 2.1** (Identité de Woodbury)

Soit  $A$  une matrice  $N \times N$ ,  $U$  une matrice  $N \times K$ ,  $B$  une matrice  $K \times K$  et  $V$  une matrice  $K \times N$ . De même, on suppose que tous les inverses matriciels considérés existent. Alors :

$$(A + UBV)^{-1} = A^{-1} - A^{-1}U \left( B^{-1} + VA^{-1}U \right)^{-1} VA^{-1}$$

En effet, si on pose  $A = \frac{1}{\sqrt{n}}X_n - zI_n$ ,  $B = M$ ,  $U = \frac{q_0}{\sqrt{n}}J$  et  $V = \frac{q_0}{\sqrt{n}}J^T$ , on obtient le résultat suivant:

$$\begin{aligned} \frac{1}{2\pi i} \oint_{\Gamma} \frac{1}{n_a} j_a^T \left( \frac{1}{\sqrt{n}}B - zI_n \right)^{-1} j_a dz &= \frac{1}{2\pi i} \oint_{\Gamma} \frac{1}{n_a} j_a^T \left( \frac{1}{\sqrt{n}}X_n - zI_n + \frac{q_0^2}{n} JMJ^T \right)^{-1} j_a dz \\ &= \frac{1}{2\pi i} \oint_{\Gamma} \frac{1}{n_a} j_a^T \left[ Q_{\frac{1}{\sqrt{n}}X_n}(z) - \right. \\ &\quad \left. \frac{q_0^2}{n} Q_{\frac{1}{\sqrt{n}}X_n}(z) J \left( M^{-1} + \frac{q_0^2}{n} J^T Q_{\frac{1}{\sqrt{n}}X_n}(z) J \right)^{-1} J^T Q_{\frac{1}{\sqrt{n}}X_n}(z) \right] j_a dz \end{aligned}$$

pour  $a \in \{1, 2\}$ .

Or, on peut remarquer que:

1.

$$\forall u^*, v, \quad u^* Q(z)v - g_{sc}(z)(q_0 \sqrt{1 - q_0^2}) u^* v \xrightarrow{n \rightarrow \infty} 0$$

D'où, on en déduit que terme à terme:  $J^T Q_{\frac{1}{\sqrt{n}}X_n}(z) J \xrightarrow{n \rightarrow \infty} g_{sc}(z)(q_0 \sqrt{1 - q_0^2}) J^T J$

$$2. \text{ Ainsi: } \left( M^{-1} + \frac{q_0^2}{n} J^T Q_{\frac{1}{\sqrt{n}}X_n}(z) J \right)^{-1} \xrightarrow{n \rightarrow \infty} \begin{pmatrix} \frac{1}{\frac{1}{M_{11}} + g_{sc}(z)(q_0^3 \sqrt{1 - q_0^2}) c_1} & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \frac{1}{\frac{1}{M_{KK}} + g_{sc}(z)(q_0^3 \sqrt{1 - q_0^2}) c_K} \end{pmatrix}$$

$$3. \text{ De plus, } j_a^T Q_{\frac{1}{\sqrt{n}}X_n}(z) J \xrightarrow{n \rightarrow \infty} g_{sc}(z)(q_0 \sqrt{1 - q_0^2}) j_a^T J \text{ et } J^T Q_{\frac{1}{\sqrt{n}}X_n}(z) j_a \xrightarrow{n \rightarrow \infty} g_{sc}(z)(q_0 \sqrt{1 - q_0^2}) J^T j_a$$

Ainsi, en passant à la limite sous l'intégrale par théorème de convergence dominé (domination classique de la résolvante par  $\frac{1}{|\Im(z)|}$ ), on obtient:

$$\begin{aligned} \frac{1}{2\pi i} \oint_{\Gamma} \frac{1}{n_a} j_a^T \left( \frac{1}{\sqrt{n}}B - zI_n \right)^{-1} j_a dz &\xrightarrow{n \rightarrow \infty} \frac{1}{2\pi i} \oint_{\Gamma} g_{sc}(z) - q_0^4(1 - q_0^2) g_{sc}^2(z) \frac{c_a}{\frac{1}{M_{aa}} + g_{sc}(z)(q_0^3 \sqrt{1 - q_0^2}) c_a} dz \\ &\xrightarrow{n \rightarrow \infty} \frac{1}{2\pi i} \oint_{\Gamma} g_{sc}(z) - q_0 \sqrt{1 - q_0^2} \frac{g_{sc}^2(z) q_0^3 \sqrt{1 - q_0^2} M_{aa} c_a}{1 + g_{sc}(z)(q_0^3 \sqrt{1 - q_0^2}) M_{aa} c_a} dz \end{aligned}$$

Or grâce aux résultats de la question précédente, on sait que  $g_{sc}(z)(q_0^3 \sqrt{1 - q_0^2}) M_{aa} c_a \xrightarrow{z \rightarrow \lambda_a} -1$  avec  $\lambda_a$  la valeur propre isolée associée à  $a$ . Dès lors  $\lambda_a$  est un pôle du second terme. Le premier n'en n'admet aucun. Ainsi, on peut utiliser le théorème des résidus comme dans le TP 1, en prenant  $\Gamma_a$  entourant seulement  $\lambda_a$ :

$$\frac{-q_0 \sqrt{1 - q_0^2}}{2\pi i} \oint_{\Gamma_a} \frac{g_{sc}^2(z) q_0^3 \sqrt{1 - q_0^2} M_{aa} c_a}{1 + g_{sc}(z)(q_0^3 \sqrt{1 - q_0^2}) M_{aa} c_a} dz = \lim_{z \rightarrow \lambda_a} -q_0 \sqrt{1 - q_0^2} (z - \lambda_a) \frac{g_{sc}^2(z) q_0^3 \sqrt{1 - q_0^2} M_{aa} c_a}{1 + g_{sc}(z)(q_0^3 \sqrt{1 - q_0^2}) M_{aa} c_a}$$

On peut ensuite utiliser la règle de l'Hôpital:

$$\begin{aligned} \frac{-q_0\sqrt{1-q_0^2}}{2\pi\mathbf{i}} \oint_{\Gamma_a} \frac{g_{sc}^2(z)q_0^3\sqrt{1-q_0^2}M_{aa}c_a}{1+g_{sc}(z)(q_0^3\sqrt{1-q_0^2})M_{aa}c_a} dz &= -q_0\sqrt{1-q_0^2}(1-g_{sc}^2(\lambda_a)) \\ &= \frac{1}{q_0^3\sqrt{1-q_0^2}(M_{aa}c_a)^2} - q_0\sqrt{1-q_0^2} \end{aligned}$$

Ainsi, on a donc:

$$\frac{1}{2\pi\mathbf{i}} \oint_{\Gamma_a} \frac{1}{n_a} j_a^T \left( \frac{1}{\sqrt{n}} B - zI_n \right)^{-1} j_a dz = \frac{1}{q_0^3\sqrt{1-q_0^2}(M_{aa}c_a)^2} - q_0\sqrt{1-q_0^2}$$

avec  $a \in \{1, 2\}$

### 2.3

Malgré nos différents essais, nous ne sommes pas parvenus à reproduire ces résultats via des simulations.

### 2.4

On peut donc effectuer une ACP sur la matrice B puis utiliser un algorithme de type *K - means* ou *EM*.

Pour évaluer cet algorithme, l'idéal serait d'avoir un réseau réel. Dès lors on pourrait calculer A. Ensuite, on pourrait faire l'hypothèse que le réseau réel suit notre modèle avec  $q_i$  constant. Enfin, on pourrait faire une recherche exhaustive sur  $q_0$  de telle manière que  $q_0$  minimise l'erreur de classification.

## 3 - Cas Hétérogène:

### 3.1

Comme exposé en 1.4 dans le cadre de la FIGURE 5, lorsque les  $q_i$  ne sont pas égaux, on a en quelque-sortie bruité la matrice  $B$  et dès lors l'algorithme précédent ne fonctionnera plus.